Sprint

Pete Sywenki
Director, Federal Regulatory Relations

Law & External Affairs
1850 M Street, NW, Suite 1100
Washington, DC 20036
Voice 202 828 7452
Fax 202 296 3469
pete.n.sywenki@mail.sprint.com

EX PARTE

January 28, 1998

Ms. Magalie Roman Salas
Secretary
Federal Communications Commission
1919 M Street, N.W. Room 222
Washington, D.C. 20554

RE: CC Docket Nos. 96-45 and 97-160

Dear Ms. Salas,

Today, I provided the attached information regarding the use of geocoded data in the Benchmark Cost Proxy Model to Robert Loube of the Universal Service Branch of the Common Carrier Bureau.

The BCPM joint sponsors request that this notice be made a part of the record in this matter. Three copies of this letter, in accordance with Section 1.1206(a)(1), are provided for this purpose. If there are any questions, please call.

Sincerely,

Pete Sywenki

Attachment

cc: Robert Loube

The Use of Geocoded Data in the BCPM with Appendix: Cases of Imperfect Geocoded Data

In the ex parte discussion filed December 11th, 1997, the BCPM Sponsors (Sprint, US WEST, Bell South) briefly discussed how geocoded data might be used or incorporated into the BCPM. In the paragraphs below, we expand on that brief discussion.

"Geocoding" in its simplest form refers to nothing more than assigning a set of latitude/longitude coordinates or V&H coordinates to a location. For example, assume Customer Smith lives at 114 Bonny Meadow Road, New Rochelle NY. Certain computer programs exist which contain the information that the 100 block of Bonny Meadow Road in New Rochelle is found at X.X° latitude and Y.Y° longitude. This information can then be used to "locate" Customer Smith on a map, and show his location relative to another customer or a fixed point such as a LEC wire center. *The assigning of the latitude & longitude to the address IS the geocoding.*

To examine how the BCPM can use geocoded data, it is worthwhile to look at how the Hatfield Model 5.0 (HM50) uses geocoded data as an example:

The HM50 itself does not use geocoded data. Geocoded points (sets of coordinates) are not used as inputs to the model, nor are they contained in any of the model's algorithms. Rather, they are used in the pre-processing stages (performed outside the model) to create "clusters", and these clusters (and associated data) form the HM50's input.

The process is straightforward: Assume there are 10 customers in a wire center, and these 10 customers have known street addresses. Assign latitude/longitude coordinates to each street address, and locate them on a map. Group together the sets of customers who are located close to each other (subject to some criteria) and those groupings form "clusters". Those "clusters" are now the input to the Hatfield Model.

In short, geocoding is used to determine which customers are located in which cluster. Once a customer has been assigned to a cluster, the geocoded data is discarded. It is important to note that this geocoded data is never used in the HM50 to determine *where* customers are located *within* clusters (some of which can be 20 square miles or more). It is only used to assign customer counts to clusters and determine the overall size of the cluster.

The BCPM can use geocoding in exactly the same way. Since the BCPM uses ultimate grids, instead of clusters, geocoding can be used to assign customers to grids.

In the BCPM's current form, customer counts are assigned to microgrids and these microgrids are aggregated into "ultimate grids", depending on certain criteria. The "ultimate grids" are the BCPM's input, just as "clusters" are the HM50's input.

Currently, customers are assigned to microgrids based on the census blocks that contain (or are contained by) those microgrids. For example, assume the Census Bureau lists 15 households in one particular census block. If that census block is covered by 2 micro grids, the 15 customers will be assigned to those 2 microgrids.

If the two microgrids contain equal amounts of road miles, the customers are split evenly between the two. If one of the microgrids contains twice as many road miles as the other, that microgrid will be assigned twice as many households as the other.[1]

It is straightforward to use geocoding to assign customers to microgrids. If the 15 customers' street addresses are known, all 15 can be given latitude/longitude or V&H coordinates. Since the grids are actually determined by latitude and longitude (recall each microgrid is 1/200th of a degree latitude/longitude), it is a simple matter to use the latitude/longitude coordinates to place the right customers in each microgrid.

An important note: In terms of actual real-time computer work, the program to do this is actually *less complex* and *more straightforward* than the program currently used to allocate customers based on road miles. It would collapse four preprocessing steps into two, and take considerably less time.

From that point, the BCPM preprocessing would continue as before. Microgrids would be aggregated into ultimate grids, based again on certain criteria, and these ultimate grids would be the units of input for the BCPM. It is possible that these ultimate grids might differ slightly from the ultimate grids the model currently uses.[2]

In summary, the BCPM in its current form is 100% capable of using geocoded data. The data would be used in the preprocessing stages to assign customers to grids, exactly as the Hatfield Model uses the data in *its* preprocessing to assign customers to clusters.

Appendix: Imperfect Geocoded Data

The BCPM, in its current form, is 100% capable of using imperfect (or partial) geocoded data as well. In the paragraphs below this approach is outlined. At this point, however, it is necessary for the BCPM Sponsors to reiterate their strong objections to this

[1] The relationship between population distribution and road distribution has been validated econometrically and this evidence has been presented to the FCC (BCPM Sponsors ex parte presentation 1/9/98). Also, it should be noted that allocation along road miles is done for all grids larger than ¼ square mile. For grids less than ¼ square mile in area, allocation is done over land area. In these extremely small grids, road miles are often evenly distributed over area so they are, in essence, one in the same.

[2] Because ultimate grids are based on certain numbers of customers within a certain distance of each other, it may be that assigning different numbers of customers to microgrids causes the microgrids to be *grouped* in a different way, forming different ultimate grids.

method. As stated in numerous ex parte presentations, the BCPM Sponsors initially considered using geocoded data in the BCPM and *rejected* this method because of the huge number of customers whose locations have not and cannot be geocoded.

That said, the BCPM could use imperfect geocoded data by adjusting the preprocessing in a manner similar to that outlined above. Initially, all known (good) geocoded locations would be allocated to microgrids using latitude/longitude coordinates.

Remaining locations would be "residual" locations. The "residuals" would equal the difference between the number of known locations in a census block (from the Census Bureau) and the number of locations geocoded in the process above for that same census block (the sum of all geocoded points in the microgrids that make up that census block). These are roughly the equivalent of Hatfield's surrogate points.

Residual points would then be geocoded (assigned latitude/longitude coordinates) as they are distributed uniformly throughout the road network of their specific census block. After the residual points were geocoded, they would be assigned to microgrids based on their coordinates just as non-residual points are.

The preprocessing would then reflect three stages: 1) allocation of good geocoded points to microgrids; 2) assignment of latitude/longitude coordinates to residual points, based on CB road networks; and 3) allocation of residual points to microgrids.

Although the BCPM Sponsors do not advocate using this approach, it is vastly superior to Hatfield's surrogate approach for two reasons. At one extreme, in the case of a census block having 0% "good" geocoded points, the result would exactly mirror the current BCPM distribution of locations. As stated earlier, our current algorithm ensures that relative locations within a grid are *maintained* (i.e. an empty southeast quadrant remains empty!). Relative locations are *discarded* under the Hatfield approach of uniform distribution within clusters.

Second, this approach (unlike the subjective Hatfield surrogate approach) allows the use of all roads in a given census block, including internal roads. Statistical tests have shown that population is highly correlated with all road miles in an area, not just peripheral road miles. To ignore internal roads distorts the accuracy of the distribution.

After all points are geocoded and assigned to microgrids, grid aggregation occurs as before (and again, ultimate grids may vary depending on this new, initial allocation). The ultimate grids serve as input to the BCPM, and the model itself (including all algorithms, equations, processes, etc.) is unaffected.